

PR #6858 完整报告

verl-project/verl

[misc] chore: fix npu parameters and typo

合并时间: 2026-06-26 16:55

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6858>

执行摘要

- 一句话: 修复 Ascend 启动脚本的参数和拼写错误
- 推荐动作: 简单维护性 PR, 无需精读。对于在 Ascend 上复现实验的用户, 建议同步更新脚本。

功能与动机

PR body 中明确说明是 fix npu parameters and typo, 解决 Ascend 环境下启动脚本的参数配置错误。

实现拆解

1. 修改 PPO 启动脚本([examples/ascend_extras/ppo_trainer/run_qwen3_8b_fsdp.sh](#)):
 - 将 data.train_batch_size 从 32 改为 64, actor_rollout_ref.actor.ppo_mini_batch_size 从 32 改为 64, 增大批量以提升吞吐。
 - 将 actor_rollout_ref.rollout.calculate_log_probs 从 True 改为 False, 禁用不必要的 log-prob 计算。
2. 修改 GRPO 启动脚本([examples/ascend_extras/grpo_trainer/run_qwen3_5_122b_a10b_32k_megatron.sh](#)):
 - 将环境变量 NDEVICES_PER_NODE 更正为 NPUS_PER_NODE, 修正拼写错误。

关键文件:

- [examples/ascend_extras/ppo_trainer/run_qwen3_8b_fsdp.sh](#) (模块 Ascend 示例; 类别 other; 类型 configuration): 修改了 batch size 和 log-prob 计算开关, 影响训练吞吐和 PPO 行为。
- [examples/ascend_extras/grpo_trainer/run_qwen3_5_122b_a10b_32k_megatron.sh](#) (模块 Ascend 示例; 类别 other; 类型 configuration): 修正变量名拼写错误, 确保 GRPO 训练器能正确读取 NPU 数量。

关键符号: 未识别

关键源码片段

[examples/ascend_extras/ppo_trainer/run_qwen3_8b_fsdp.sh](#)

修改了 batch size 和 log-prob 计算开关，影响训练吞吐和 PPO 行为。

```
# 在 run_qwen3_8b_fsdp.sh 中优化批量大小并禁用 log-prob 计算
# 增大批量大小以提升 NPU 利用率
data.train_batch_size=64
actor_rollout_ref.actor.ppo_mini_batch_size=64
# 禁用 log-prob 计算（Ascend 场景下非必要）
actor_rollout_ref.rollout.calculate_log_probs=False
```

评论区精华

无人工 review 讨论，仅由 Gemini Code Assist 自动评论总结变更，以及由 wucong25 直接批准。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低，仅涉及 shell 脚本中的参数配置变更，无核心代码逻辑改动。
calculate_log_probs=False 可能影响依赖该值的后续计算（如 PPO 算法的 advantage 估计），但原脚本本应如此配置以匹配 Ascend 性能优化。
- 影响：仅影响 Ascend 环境下使用该启动脚本的用户。PPO 脚本参数优化可提升训练吞吐。
GRPO 脚本变量名修正可防止启动失败。
- 风险标记：配置变更

关联脉络

- PR #6807 [model] feat: add qwen3-122b long seq launch script for ascend: 同为 Ascend 启动脚本变更，修改了同一个 GRPO 脚本文件。